No. 17-1625

In the Supreme Court of the United States

RIMINI STREET, INC., AND SETH RAVIN,

Petitioners,

v.

ORACLE USA, INC., ORACLE AMERICA, INC., AND ORACLE INTERNATIONAL CORPORATION.

On Writ of Certiorari to the United States Court of Appeals for the Ninth Circuit

BRIEF OF *AMICI CURIAE* SCHOLARS OF CORPUS LINGUISTICS SUPPORTING PETITIONERS

JAMES A. HEILPERN
Law & Corpus Linguistics Fellow
BYU LAW SCHOOL
415 JRCB
Provo, UT 84602

GENE C. SCHAERR
Counsel of Record
ERIK S. JAFFE
MICHAEL T. WORLEY
SCHAERR | JAFFE LLP
1717 K Street NW, Suite 900
Washington, DC 20006
(202) 787-1060
gschaerr@schaerr-jaffe.com

QUESTION PRESENTED

Whether the Copyright Act's allowance of "*full costs*" (17 U.S.C. 505) is limited to the categories and amounts of costs taxable under 28 U.S.C. 1920 & 1821, or also authorizes an award of expert witness fees and other "non-taxable" expenses.

TABLE OF CONTENTS

Page

QUESTION PRESENTED i

TABLE OF CONTENTS ii

TABLE OF AUTHORITIES iii

INTRODUCTION AND INTERESTS OF *AMICI* 1

STATEMENT 3

SUMMARY OF ARGUMENT 5

ARGUMENT 7

I. The Court should consider evidence derived from corpus linguistics to test the Ninth Circuit's interpretation of the Copyright Act. 7

II. Any meaning the Court attributes to "full" must be limited and controlled by the term of art it modifies. 19

III. The word "full" is often delexicalized. 24

A. "Delexicalization" is common in American English 25

B. "Full" is often delexicalized in ordinary American English 26

C. "Full" is delexicalized in legalese. 30

IV. Judges and Congress consistently use the phrase "full costs" in a way that excludes expert witness fees. 32

CONCLUSION 38

TABLE OF AUTHORITIES

Page(s)

Cases

Arlington Central School District v. *Murphy*,
548 U.S. 291 (2006) 3, 7

Bank of United States v. *Bank of Wash.*,
31 U.S. 8 (1832) 30

Brooks v. *Marbury*,
24 U.S. 78 (1826) 30

Chambers v. *NASCO, Inc.*,
501 U.S. 32 (1991) 23

Crawford Fitting Co. v. *J.T. Gibbons, Inc.*,
482 U.S. 437 (1987) 3, 7

Dept. of Housing & Urban Develop. v. *Rucker*,
535 U.S. 125 (2002) 19

District of Columbia v. *Heller*,
554 U.S. 570 (2008) 9, 10

Express LLC v. *Forever 21*,
2010 WL 11512410 (C.D. Cal. Nov. 15, 2010) 34

FCC v. *AT&T, Inc.*,
562 U.S. 397 (2011) 11

Fire Ins. Exchange v. *Oltmanns*,
416 P.3d 1148 (Utah 2017) 12

Hively v. *Ivy Tech Community College of Indiana*,
853 F.3d 339 (7th Cir. 2017) 15

Hygrade No. 24, the v. *the Dynamic*,
143 F. Supp. 634 (E.D.N.Y. 1955) 35

Imaginary Images Inc. v. *Evans*,
2009 WL 2488004
(E.D. Va. Aug. 12, 2009) 34

In re Adoption of Baby E.Z.,
266 P.3d 702 (Utah 2011) 11

In re Lipset,
119 F. 379 (S.D.N.Y. 1902) 34

Innovative Sports Management, Inc. v. *Adriana, Inc.*,
2014 WL 1491339 (E.D. Va. Feb. 3, 2014) 34

J&J Sports Prods., Inc. v. *Evolution Entertainment Grp., LLC*,
2014 WL 2066809 (E.D. La. 2014) 34

Lamore v. *Ives*,
977 F.2d 713 (1st Cir. 1992) 14

Marx v. *General Revenue Group*,
568 U.S. 371 (2013) 7

McBoyle v. *United States*,
238 U.S. 25 (1931) 14

MCI v. *AT&T*,
512 U.S. 218 (1994) 15, 16

Mich. Am. Fed. Of State Cty. & Mun. Employees Council 25, Local 1640 v. *Matrix Hum. Servs.*,
589 F.3d 851 (6th Cir. 2009) 14

Muscarello v. *United States*,
524 U.S. 125 (1998) 10

My Pie Int'l, Inc. v. *Debould, Inc.*,
687 F.2d 919 (7th Cir. 1982) 14

Ogden v. *Saunders*,
25 U.S. 213 (1827) 30

People v. *Harris*,
885 N.W.2d 832 (Mich. 2016) 12

Ramsay v. *Allegre*,
25 U.S. 611 (1827) 31

Randall v. *Loftsgaarden*,
478 U.S. 647 (1986) 16

Rowland v. *Cal. Men's Colony*,
506 U.S. 194 (1993) 19

Samsung Electonics v. *Apple*,
137 S. Ct. 429 (2016) 16

SEC v. *Utsick*,
2009 WL 1606511 35

State v. *Canton*,
308 P.3d 517 (Utah 2013) 12

State v. *Rasabout*,
356 P.3d 1258 (Utah 2015) 10, 15

Talamini v. *Allstate Ins. Co.*,
470 U.S. 1067 (1985) 24

The Betsy,
3 F. Cas. 299 (D. Mass. 1815) 35

Twentieth Century Fox Film Corp. v. *Entertainment Distributing*,
429 F.3d 869 (9th Cir. 2005) 3, 4, 8

United States v. *Am. Trucking Ass'ns, Inc.*,
310 U.S. 534 (1940) 14

United States v. Costello,
666 F.3d 1040 (7th Cir. 2012) 10

United States v. *Goldenberg*,
168 U.S. 95 (1897) 19

United States v. *Reese*,
92 U.S. 214 (1875) 34

Waste Control Specialists, LLC v. *Envirocare of Tex., Inc.*,
199 F.3d 781 (5th Cir. 2000) 34

West Virginia University Hospitals, Inc. v. *Casey*,
499 U.S. 83 (1991) 33

Statutes

5 U.S.C. 3330c 37

7 U.S.C. 26 37

12 U.S.C. 5567 36

15 U.S.C. 2060 36

17 U.S.C. 505 ..... i, 2, 3, 5, 6, 18, 22, 27, 30, 32

30 U.S.C. 1270 37

30 U.S.C. 1427 36

31 U.S.C. 3554 36

42 U.S.C. 9607 36

49 U.S.C. 30171 36

Other Authorities

A. Spencer, et. al., *The Oxford Handbook of Compounding*,
47 J. Linguistics 481 (2011) 20

Abbe R. Gluck & Lisa Schultz Bressman, *Statutory Interpretation From the Inside—An Empirical Study of Congressional Drafting, Delegation, and the Canons: Part I*,
65 Stan. L. Rev. 901, 934 (2013) 32

Antonin Scalia & Bryan A. Garner, Reading Law (2012) 19, 24, 25, 30, 32

Brief for Open Government Project as Amicus Curiae, FCC v. AT&T, 562 U.S. 397 (2011) 13

Corpus of Contemporary American English 12, 17, 18, 22, 28

Corpus of Historical American English .. 17, 18, 22, 27

Corpus of the Supreme Court of the United States 17, 22

Douglas Biber, *et. al.*, Corpus Linguistics: Investigating Language Structure and Use (1998) 10

Elena Tognini-Bonelli, Corpus Linguistics at Work 116 (2001) 26

H.R. Howard, The History of Virgil A. Stewart and his Adventure in capturing and Exposing the Great 'Western Pirate' and his Gang 165 (1836)..... 27

Henry Charles Carey, Essay on the Rate of Wages, with an Examination of the Causses of the Differences in the Condition of the Labouring Population Throughout the World 232 (1835)..... 28

James C. Phillips & Jesse Egbert, *Advancing Law & Corpus Linguistics*, 2017 B.Y.U. L. Rev. 1589 9

James Cleith Phillips & Sara White, *The Meaning of the Three Emoluments Clauses in the U.S.* Constitution, 59 S. Tex. L. Rev. 181 (2018..... 13

James Heilpern, *Dialects of Art*, 58 Jurimetrics 377, 380 (2018)..... 14

John Sinclair, Corpus, Concordance, Collocation (1991)..... 24

John Sinclair, Trust the Text: Language, Corpus and Discourse 20 (2004) 24, 25, 26

Kiel Brennan-Marquez & Jill Anderson, *The "Lawyer v. Human" Problem in Corpus Linguistics*, ... 15

Lawrence B. Solum, *The Fixation Thesis: The Role of Historical Fact in Original Meaning*, 91 Notre Dame L. Rev. 1 (2015) 16

Lawrence M. Solan & Tammy Gales, *Corpus Linguistics as a Tool in Legal Interpretation*, 2017 B.Y.U. L. Rev. 1311 31

Lawrence Solan, *et. al.*, *False Consensus Bias in Contract Interpretation*, 108 Colum. L. Rev. 1268 (2001) 10

Linda D. Jellum, Mastering Statutory Interpretation 104 (2008) 31, 32

Martin Haspelmath, *On understanding word order asymmetries*, 28 Theoretical Ling. 159 (2002) 20

Robert Montgomery Bird, The Infidel's Doom 109 (1840) 28

Stephen C. Mouritsen, *The Dictionary is not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning*, 2010 B.Y.U L. Rev. 1915, 1954–1955 8

Stephen Mouritsen, *Contract Interpretation with Corpus Linguistics*, 94 Wash. L. Rev. (forthcoming 2019) 24, 25

Thomas R. Lee & Stephen C. Mouritsen, *Judging Ordinary Meaning*, 127 Yale L.J. (forthcoming 2018). 1, 8, 11

V. Gonzalez-Diaz, *Great Big Stories and Tiny Little Changes: Tautological size-adjective clusters in present-day English*, 22 Eng. Lang. & Ling. 499 (2018) 26

William Baude & Stephen E. Sachs, *The Law of Interpretation*, 130 Harv. L. Rev. 1079, 1125 (2017) 24

William Jay, *The Life of John Jay*, 37 N. Am. Rev. 315, 316 (1833) 28

William Sullivan & George Barrell Emerson, The Political Class Book: Intended to Instruct the Higher Classes in Schools 120 (1831) 28

INTRODUCTION AND INTERESTS OF *AMICI*[1]

Amici (listed in the Appendix) are scholars of a methodology for answering questions of interpretation in a systematic, rigorous manner—a methodology known as "corpus linguistics." As Utah Supreme Court Justice Thomas R. Lee and Stephen Mouritsen put it, "corpus linguistics is an empirical approach to the study of language that involves large, electronic databases," which are used to "draw inferences about language from data gleaned from real-world language in its natural habitat—in books, magazines, newspapers, and even transcripts of spoken language."[2] Because judges—like linguists and lexicographers—are interested in the "original public meaning" of historic texts and the "ordinary meaning" of modern texts, *amici* believe these databases can be invaluable in resolving difficult questions of constitutional and statutory interpretation.

Usage evidence derived from these databases demonstrates that the Ninth Circuit's holding is linguistically untenable. One reason is that an adjective's meaning is generally derived from the noun it modifies, not the other way around: In this case, "full" can no more alter the meaning of "costs" than it can the meaning of "moon," "speed," "time," "parking lot," or "house." In fact the noun and its context tells us which of the many meanings of "full" was intended.

---

[1] With the exception of BYU Law School—which has generously covered the costs of printing this brief—no one other than *amici* and their counsel authored any part of this brief or made a monetary contribution to fund its preparation or submission. Petitioners and Respondents both filed blanket consents.

[2] Thomas R. Lee & Stephen C. Mouritsen, 127 Yale L.J. 788, 827 (2017).

Here there is also strong evidence that "full" in Section 505 should be considered a "delexicalized" adjective. A delexicalized adjective is one whose purpose is to draw attention to and underline an attribute that is already embedded in the meaning of the noun such as "clenched fist" and "free gift." "Full" often serves to emphasize the completeness of an object that is already presumed to be complete—like "full deck of cards," "full set of teeth," and "full costs." As applied here, then, "full costs" merely means all the costs that are *otherwise* authorized by the relevant law—not all costs that might be imagined.

STATEMENT

Petitioner Rimini Street, Inc. is a company that provides after-market support services to Respondent Oracle's software clients. Pet. App. 5a. The lower courts found that the *way* Rimini Street provided these services infringed Oracle's software copyrights. Whether Rimini Street actually violated Oracle's copyright is not on appeal. Instead, this case focuses on the Ninth Circuit's interpretation of the Copyright Act's costs provision.

In addition to damages, the Copyright Act allows a court to award a prevailing party "full costs." 17 U.S.C. 505. This Court has held that the word "costs" by itself is a term of art which refers to just six discrete categories of litigation expenses enumerated in 28 U.S.C. 1920 & 1821: fees for the clerk and marshal; transcript fees; disbursements for printing and witnesses; fees for making copies; docketing fees; and the compensation of court-appointed experts and certain special interpretation services. *Arlington Central School District* v. *Murphy*, 548 U.S. 291 (2006). Other litigation costs—such as expert witness fees—are not recoverable in the absence of "plain evidence" of "clear ... congressional intent to supersede those sections." *Crawford Fitting Co.* v. *J.T. Gibbons, Inc.*, 482 U.S. 437, 445 (1987).

Thirteen years ago, the Ninth Circuit concluded that the Copyright Act allowed courts to award the full panoply of litigation expenses as part of "costs." *Twentieth Century Fox Film Corp.* v. *Entertainment Distributing*, 429 F.3d 869, 885 (9th Cir. 2005). "[T]o do otherwise would"—according to the Ninth Circuit—"violate the long standing principle of statute interpretation that statutes should be construed to make

surplusage of any provision." *Id.* After all, any other reading of the Copyright Act would supposedly "read[] the word 'full' out of the statute." *Id.* Bound by this precedent, the district court below held that "full costs" meant "all costs incurred in litigation." JA305. It therefore ordered Rimini Street to reimburse Oracle for over $12 million in litigation expenses not typically considered "costs"—including (among other things) expert witness fees, additional e-discovery fees, contract attorney services, and jury consulting. *Id.* Rimini Street appealed to the Ninth Circuit, which affirmed on *stare decisis* grounds. JA346. Rimini Street sought rehearing en banc, which was denied. This Court granted certiorari.

SUMMARY OF ARGUMENT

Petitioners explain that the term "full costs" as used in the Copyright Act's costs provision should be limited to those categories of litigation expenses specifically enumerated by Congress in 28 U.S.C. 1920 and 1821. This reading is strongly supported by evidence derived from a technique called corpus linguistics—the use of electronically-searchable linguistic databases to investigate the meaning and function of words within a particular community at a particular time. Application of that methodology shows that the Ninth Circuit's holding is linguistically untenable.

As explained above and in Section II below, an adjective's meaning is "disambiguated"—that is, clarified or made less ambiguous—by the noun it modifies, not vice versa. Thus, "full" can no more alter the meaning of "costs" than it can the meaning of "moon," "speed," "time," "parking lot," or "house." Instead, the noun associated with "full," and the context in which the phrase is used, indicates which of the many meanings of "full" was intended.

Here, the linguistic evidence shows that the "full" in Section 505 should be considered a "delexicalized" adjective, that is, an adjective whose purpose is to draw attention to and underline an attribute already fundamental to the nature of the noun that is already embedded in the meaning of the noun. "Full" often serves to emphasize the completeness of an object that is already presumed to be complete, like "full deck of cards," "full set of teeth," and "full costs."

These findings are supported by the linguistic conventions of Congress and the courts, which frequently use the terms "full costs" and "expert witness fees" in

a way that makes clear that the latter is not an element of the former.

ARGUMENT

I. The Court should consider evidence derived from corpus linguistics to test the Ninth Circuit's interpretation of the Copyright Act.

The Copyright Act states that a "court in its discretion may allow the recovery of full costs by or against any party other than the United States." 17 U.S.C. 505 (emphasis added). This Court has already stated the word "costs" is a "term of art that does not generally include expert fees." *Arlington Central School District* v. *Murphy*, 548 U.S. 291 (2006). Instead, it is limited to the six discrete categories enumerated in 28 U.S.C. 1920 and 1821. See, *e.g.*, *id.*; *Tanguchi* v. *Kan Pacific Saipan, Ltd.*, 566 U.S. 560, 562 (2012). As this Court has put it, "[a]ny argument that a federal court is empowered to exceed the limitations explicitly set out in §§ 1920 and 1821" when awarding costs must fail unless there is "plain evidence" of "*clear* ... congressional intent to supersede those sections." *Crawford Fitting Co.* v. *J.T. Gibbons, Inc.*, 482 U.S. 437, 445 (1987). As always, "[t]he best evidence of congressional intent ... is the statutory text that Congress enacted." *Marx* v. *General Revenue Group*, 568 U.S. 371, 392 n. 4 (2013) (Sotomayor, J. dissenting) (citing *W. Va. Univ. Hosp., Inc.* v. *Casey*, 499 U.S. 83, 98 (1991)).

The courts below claim to have found this evidence in the phrase "full costs." Based on Ninth Circuit precedent, the district court interpreted the phrase "full costs" to mean "all costs incurred in litigation." *Id.* The Ninth Circuit had not only concluded that this was the best reading of the statute, but thought "there [could] be no other import to the phrase" because any other interpretation would "effectively read[] the word 'full'

out of the statute." *Twentieth Century Fox Film Corp.* v. *Entertainment Distributing*, 429 F.3d 869, 885 (2005). This is a testable hypothesis, one that can be assessed by analytical tools and techniques developed by linguists and lexicographers over the last fifty years. This methodology—which has only recently gained traction in the law—is known as corpus linguistics.

1. Corpus linguistics is a discipline which investigates real-language use and function by analyzing electronic databases of naturally-occurring texts. Thomas R. Lee & Stephen C. Mouritsen, *Judging Ordinary Meaning*, 127 Yale L.J. 788 (2018). These databases have an esoteric name—corpora (the plural of corpus)—but are simply digitally-searchable collections of real-world sources: books, newspapers, speeches, scholarly articles, television transcripts, etc. *Id.* at 33. The sources are said to occur "naturally" because they "were not elicited for the purpose of the study. That is ... no one ask[ed] the speakers or writers whose words are represented in the corpus to speak or write for the purpose of subjecting their words to linguistic scrutiny. Instead, the architect of the corpus assemble[d] her collection of speech and writing samples after the fact." Stephen C. Mouritsen, *The Dictionary is not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning*, 2010 B.Y.U L. Rev. 1915, 1954–1955.

Although corpus linguistics offers many different tools, when interpreting a historical text such as a statute, the following method is helpful:

- Identify a corpus that corresponds with and is representative of the speech community and time period she wishes to investigate;

- Search for the relevant search term using a "Keyword in Context" (KWIC) feature, which finds and displays in context every instance of the queried term in the database;
- Generate a random (and thus likely representative) sample of the returned KWIC lines large enough to detect significant effects statistically; and
- Code each KWIC line in the sample for its relevant word sense, relying on the system's expanded context feature when necessary.[3]

This approach is similar to that used by many lexicographers today. It can help produce useful quantitative and qualitative evidence about the real-world usage of the relevant term.

Of course, not all of the examples produced by the corpus will be helpful. Like Google or Westlaw, a corpus search will sometimes identify sources in which the queried term was used in a very different context. Other times the usage will be vague or ambiguous. But analyzing a random sample of concordance lines, as a whole, produces a broader picture of language usage, defining the range of potential meanings a word or phrase may take and often revealing trends and patterns that otherwise would have remained unnoticed.

2. This approach should feel familiar to most judges. Courts search for real-world examples of linguistic usage to help make sense of legal passages all the time. For example, in *Heller*, opinions on both sides

---

[3] See, *e.g.*, James Cleith Phillips & Jesse Egbert, *Advancing Law & Corpus Linguistics*, 2018 B.Y.U. L. Rev. (forthcoming), *available at* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3057415.

of the issue considered concrete examples mined from "founding era sources" of how the phrases "keep arms" and "bear arms" were actually used by contemporaries of the Constitution. 554 U.S. at 581-592.

The problem is that until now "a judge has [had] no way of determining whether she is correct in her assessment that her own interpretation is widely shared." Lawrence Solan, et al., *False Consensus Bias in Contract Interpretation*, 108 Colum. L. Rev. 1268, 1273 (2001). Linguists have long noted the fallibility of human linguistic intuition. Because "humans tend to notice unusual occurrences [of words] more than typical ones," judges run the risk of over-crediting the frequency of obscure word senses. Douglas Biber et al., Corpus Linguistics: Investigating Language Structure and Use 3 (1998).

To combat this, some judges have turned to electronic databases "in an effort to assemble a greater number of examples than ... can [be] summon[ed] by memory" alone, *State* v. *Rasabout*, 356 P.3d 1258, 1271 (Utah 2015) (Lee, A.C.J.), as a way to check their linguistic intuition. This Court did so in *Muscarello* v. *United States*, "survey[ing] modern press usage [of the word "carry"] by searching computerized newspaper databases." 524 U.S. 125, 129-130 (1998). Likewise, in *United States* v. *Costello*, Judge Posner performed a Google search "of several terms in which the word 'harboring' appears" on the "supposition that the number of hits per term is a rough index of the frequency of its use." 666 F.3d 1040, 1044 (7th Cir. 2012). While these approaches had some methodological shortcomings,[4]

---

[4] See, *e.g.*, *Muscarello*, 524 U.S. at 129 (Breyer, J.) (describing his own approach as "crude[]"); *State* v. *Rasabout*, 356 P.3d 1258,

they are laudable for their efforts to check the court's linguistic assumptions.

Corpus analysis simply empowers the judge to do this kind of linguistic research in a systematic fashion. Subsequent researchers can test the validity of any dataset by performing the corpus search again on their own. This approach does not supplant the judge as the ultimate decision maker. It simply furnishes the judge with more and better evidence to help inform her ultimate decision while simultaneously making the decision-making process more transparent.

3. Over the last few years, some judges have cautiously begun applying corpus linguistic tools and techniques to help resolve difficult cases. For example, in 2011 Justice Ginsburg cited corpus linguistics evidence during oral arguments in *FCC* v. *AT&T, Inc.*, 562 U.S. 397 (2011). See Transcript of Oral Argument in No. 09-1279 at 37. The case boiled down to whether the word "personal" as used in the Freedom of Information Act was merely the 'adjectival form' "of the noun person" so that the phrase "personal privacy" encompassed corporate privacy. See 562 U.S. at 406. While Chief Justice Roberts' majority opinion did not cite corpus linguistics directly, its reasoning largely tracked the amicus brief by the Project on Government Oversight, which did.

That same year, Justice Lee of the Utah Supreme Court became the first judge in the country to expressly use corpus linguistics in an opinion. *In re Adoption of Baby E.Z.*, 266 P.3d 702 (Utah 2011) (Lee,

---

1280 (Utah 2015) (Lee, J. concurring) (critiquing Judge Posner's reliance on Google searches); Lee & Mouritsen, *supra* at 812-813 (same).

J. concurring). Relying on evidence drawn from Brigham Young University's Corpus of Contemporary American English (COCA),[5] he concluded that the term "custody determination" as used in the federal Parental Kidnapping Prevention Act did not extend to adoption proceedings because "the most common family-law sense of the word 'custody' occurs in the setting of a divorce." *Id.* at 724.

Since then, a number of other judges and justices around the country have followed suit.[6] For example, in the Michigan Supreme Court case *People* v. *Harris*, both the majority and dissent relied on the COCA to analyze whether someone had been forced to make an involuntary statement if the "information" he provided law enforcement officers was actually false. 885 N.W.2d 832. Justice Thomas likewise used corpus linguistics at the end of last Term in his dissent in *Carpenter* v. *United States*.[7]

---

[5] Corpus of Contemporary American English, https://corpus.byu.edu/coca/.

[6] See, *e.g.*, *American Bankers Ass'n* v. *Nat'l Cred. Union Admin.*, 306 F. Supp. 44 (D.D.C. 2018) (citing the Corpus of Historical American English); *Fire Ins. Exchange* v. *Oltmanns*, 416 P.3d 1148, 1163 n.9 (Utah 2017) (Durham, J. concurring) ("[Corpus linguistic] tools for empirical analysis are readily available for lawyers and should be used when appropriate"); cf. *State* v. *Canton*, 308 P.3d 517 (interpreting the phrase "out of the state" based on an analysis of the use of that phrase in newspaper articles compiled through a Google News search).

[7] *Carpenter* v. *United States*, 138 S. Ct. 2206, 2238 nn. 2-5 (2018) (Thomas, J. dissenting) (citing four electronic databases of early American texts including the Corpus of Historical American English). Justice Thomas also cited a law review article that relied heavily on corpus linguistics in his concurring opinion in *Lucia* v. *S.E.C.*, 138 S. Ct. 2044, 2056-57 (Thomas, J. concurring) (citing

4. Before undertaking a corpus analysis, a judge must make certain preliminary decisions—namely, she must decide what the relevant speech community and time period should be. Evidence gleaned from a corpus is only helpful if the sources contained in the corpus are representative of the relevant speech community. James Cleith Phillips & Sara White, *The Meaning of the Three Emoluments Clauses in the U.S.* Constitution, 59 S. Tex. L. Rev. 181 (2018). A corpus composed entirely of transcripts from Argentinian telenovelas—no matter how large or meticulously scrutinized—cannot provide relevant evidence for investigating the speech patterns of American diplomats. Likewise, a 14th century legal corpus will likely be unhelpful in clarifying the meaning of modern statutes. But a corpus that is representative of the questions being asked is "is like Lexis on steroids." Brief for Open Government Project as Amicus Curiae at 14, *FCC* v. *AT&T Inc.*, 562 U.S. 397 (2011). Searching such a corpus for a particular term produces a set of real-world examples (called "KWIC" or "concordance lines") drawn from the database and showing how that term has actually been used within the relevant community during the relevant time period.

Unfortunately, courts are not always consistent in their answers to these preliminary questions with respect to statutory interpretation. For example, judges often justify their invocation of the ordinary meaning canon by citing the principle of fair notice. As Justice Oliver Wendell Holmes put it, "it is reasonable that a fair warning should be given to the world *in language that the common world will understand*, of what the

---

Jennifer Mascott, *Who Are "Officers of the United States"?*, 70 Stan. L. Rev. 443 (2018).

law intends to do if a certain line is crossed." *McBoyle* v. *United States*, 238 U.S. 25, 27 (1931). This would suggest that the relevant speech community for understanding any statute is whatever community is regulated by that statute. See, *e.g.*, James Heilpern, *Dialects of Art*, 58 Jurimetrics 377, 380 (2018). Generally applicable statutes would therefore be presumed to be written in the language of the average American, and the proper corpus for analyzing that language would be a large, representative database of texts written in ordinary, American English. *Id.* at 394.[8]

But judges sometimes also acknowledge that statutory language can be a "virtually impenetrable thicket of legalese and gobbledygook." *Lamore* v. *Ives*, 977 F.2d 713 (1st Cir. 1992).[9] Some judges also believe that "[i]n the interpretation of statutes, the function of the courts is ... to construe the language so as to give effect to the intent of Congress," *United States* v. *Am. Trucking Ass'ns, Inc.*, 310 U.S. 534, 542 (1940), a Congress that is demographically wealthier, better educated, and with a greater percentage of lawyers than the nation as a whole.[10] Such statements would sug-

---

[8] All of the judges that have cited corpus linguistics in opinions thus far have taken this approach.

[9] See also, *e.g. My Pie Int'l, Inc.* v. *Debould, Inc.*, 687 F.2d 919, 936 (7th Cir. 1982) (Eschbach, J. concurring in part and dissenting in part) (acknowledging that the statute was written in "nineteenth century legalese"); *Mich. Am. Fed. Of State Cty. & Mun. Employees Council 25, Local 1640* v. *Matrix Hum. Servs.*, 589 F.3d 851, 858 (6th Cir. 2009) (noting that the statute was "drafted in the 1930s using the legalese that was the vernacular of that era);

[10] Cf. Phillip Bump, *The New Congress is 80% white, 80 percent male and 92 percent Christian*, Washington Post, Jan. 5, 2015,

gest that the relevant speech community is either Congress or perhaps reasonably, well-educated lawyers, and that statutes are best understood as being written in a distinct dialect-of-art we might call *legalese*.[11] Corpus linguistics is still helpful for a judge who accepts this proposition—she would simply need to turn to a corpus of legal documents to investigate the statutory language.[12]

Professors Kiel Brennan-Marquez and Jill Anderson recently put it another way—judges disagree about whether statutes should be understood as "written-by humans" or "written-by-lawyers."[13] The answer to that question will determine the appropriate corpus.

Judges have likewise sent mixed signals about whether there is a temporal component to statutory interpretation. This Court has stated that "the most relevant time for determining a statutory term's

---

available at https://www.washingtonpost.com/news/the-fix/wp/2015/01/05/the-new-congress-is-80-percent-white-80-percent-male-and-92-percent-christian/?utm_term=.c31a1c3edf0b.

[11] Some scholars have suggested the same with respect to the Constitution. See John O. McGinnish & Michael B. Rappaport, *Original Methods Originalism: A New Theory of Originalism and the Case Against Construction*, 103 N.W. Univ. L. Rev. 751 (2009).

[12] Cf. *State* v. *Rasabout*, 359 P.3d 1258,1266 (Utah 2015) (criticizing Justice Lee's use of a corpus that lacked legislative documents because "the 'idiosyncrasies' of the Utah Legislature constitute the rule of law in this state.")

[13] Kiel Brennan-Marquez & Jill Anderson, *The "Lawyer v. Human" Problem in Corpus Linguistics*, *available at* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3287336.

meaning" is the time period in which it "became law."[14] *MCI* v. *AT&T*, 512 U.S. 218 (1994). Yet, judges also frequently rely exclusively on *modern* dictionaries to interpret historic statutes.[15] This practice seems predicated on either an assumption that statutory language never suffers from "linguistic drift,"[16] or a belief that judges—when interpreting a historic statute—should give "fresh meaning to a statement" to reflect "modern needs and understandings" when "[t]imes [and language] have changed." *Hively* v. *Ivy Tech Community College of Indiana*, 853 F.3d 339, 353-54 (7th Cir. 2017) (Posner, J., concurring).

5. Regardless of how the Court chooses to answer these preliminary questions, it would benefit from the use of usage evidence derived from corpus tools and techniques. For this reason, we have analyzed the relevant statutory terms using a variety of corpora in order to investigate the statute's meaning in both ordinary American English and legalese at the time

---

[14] Professor Lawrence Solum calls this the "Fixation Thesis." Lawrence B. Solum, *The Fixation Thesis: The Role of Historical Fact in Original Meaning*, 91 Notre Dame L. Rev. 1 (2015).

[15] See, *e.g.*, *Samsung Electonics* v. *Apple*, 137 S. Ct. 429 (2016) (relying predominantly on the American Heritage Dictionary (5th ed. 2011) to interpret the Patent Act of 1952); *Randall* v. *Loftsgaarden*, 478 U.S. 647 (1986) (Brennan, J. dissenting) (relying on Webster's Ninth New Collegiate Dictionary (1983) to interpret the Securities Act of 1933).

[16] "Linguistic Drift" refers to "the notion that language usage and meaning shifts over time." Thomas R. Lee & James Cleith Phillips, *Data-Driven Originalism*, 167 Penn. L. Rev. (forthcoming), *available at* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3036206.

the relevant provisions were passed and today. To do this, we utilized the following corpora:

- The Corpus of Contemporary American English (COCA), which "contains more that 560 million words of text (20 million words each year 1990-2017) ... equally divided among spoken, fiction, popular magazines, newspapers, and academic texts."[17]
- The Corpus of Historical American English (COHA), which "contains more than 400 million words of text from the 1810s-2000s" drawn from historical fiction books, magazines, newspapers, and non-fiction books.[18]
- The Corpus of the Supreme Court of the United States, which "includes all opinions in the United States Reports and opinions published by the Supreme Court through the 2017 term."[19]

COCA and COHA were used to investigate what we call "Ordinary American English." We limited our COHA search to only those texts within the corpus published between 1826 and 1836 in order to generate a sample of written vernacular around the time the relevant language was first passed by Congress in 1831. The Corpus of the Supreme Court—which we

---

[17] Corpus of Contemporary American English, https://corpus.byu.edu/coca/.

[18] Corpus of Historical American English, https://corpus.byu.edu/coha/.

[19] Corpus of the Supreme Court of the United States, https://lawcorpus.byu.edu/coscotus.

also limited to only those cases issued between 1826 and 1836—was used as a proxy for historic legalese.[20]

In addition, we created two small corpora of our own to supplement our findings. The first is a corpus composed of all publicly-available text on Oracle's own website. The second contains 222 contracts to which Oracle was a party, each drawn from the LawInsider Database. These were examined using AntConc, a free corpus software platform for analyzing custom corpora.

6. As previously mentioned, the Ninth Circuit's holding rests on the presumption that its reading of the Copyright Act's cost provision is the *only* plausible interpretation that gives effect to each and every word of the statute. Our analysis proves this to be false. Based upon our analysis of a simple random sample[21] of concordance lines—actual examples of usage—containing the word "full" generated from each of our five corpora, there are at least two other linguistically-sound constructions of Section 505 that do not "read full out of the statute." The most natural reading is to

---

[20] Although all of the evidence points in the same direction here, it will not always in every case. We therefore urge the Court to clearly articulate its normative preferences to the questions posed above, and then rely on the corresponding set of data provided to answer the question. (i.e. A judge who believe statutes should be interpreted in light of the vernacular of the period it was passed should rely on the COHA dataset).

[21] Each of our samples was large enough to give us at least a 95% confidence interval with only a 4% margin of error. "Full" appeared in the COHA (1826-1836), COCA, Oracle Website, and Oracle Contract corpora 4106; 11,335; 1618; and 649 times, respectively. Our samples for each were therefore 525, 600, 384, and 312. We reviewed all 1026 hits for "full" in the Corpus of the Supreme Court (1826-1836). Our raw data can be accessed at https://goo.gl/Tmjnra.

assume that "full" simply clarifies that a court may award all of the costs allowed under Section 1920 but no more. In the alternative, the Court could conclude that "full" is a delexicalized adjective (as it usually is in legal writing), meaning its primary effect is to emphasize rather than add additional semantic content to the sentence. Although both of these possibilities will be explained in greater detail below, it does not matter which reading the Court ultimately decides to adopt—the mere presence of plausible alternatives destroys the "plain evidence" of "*clear* ... congressional intent to supersede" Section 1920's statutory definition of costs that the Ninth Circuit claims to have found.

II. Any meaning the Court attributes to "full" must be limited and controlled by the term of art it modifies.

This Court has long held that its statutory interpretations must be governed by basic "rules of grammar." See *Dept. of Housing & Urban Develop.* v. *Rucker*, 535 U.S. 125, 131 (2002) (reversing lower court's interpretation because it "runs counter to basic rules of grammar"); *Rowland* v. *Cal. Men's Colony*, 506 U.S. 194, 205 (1993) (rejecting petitioner's statutory argument because "it would wrench the rules of grammar to read" the statute that way); *United States* v. *Goldenberg*, 168 U.S. 95, 103 (1897) ("[The lawmaker] is presumed to know the ... rules of grammar."); see also Scalia & Garner, Reading Law at 140 ("Words are to be given the meaning the proper grammar and usage would assign them."). The Ninth Circuit's holding violated that principle by completely ignoring the essential relationship between nouns and adjectives.

An adjective is a "word[] that modif[ies] nouns and pronouns, primarily by describing a particular quality of the word it is modifying." Adjective, dictionary.com (last visited Nov. 19, 2018). "[I]n syntactic representations the adjective is a subordinate category, a dependent of the noun ... that it modifies." A. Spencer, et al., *The Oxford Handbook of Compounding*, 47 J. Linguistics 481, 489 (2011). In layman's terms, this means that in the relationship between adjectives and their nouns, the noun is king—an adjective's meaning and scope is always relative to the noun it is modifying. The expression "a tall seven-year-old" evokes a very different standard of tallness than a "a tall NBA player."

When adjectives are polysemous—linguist-speak for having multiple but related meanings—they "depend on nouns for ambiguity resolution." Martin Haspelmath, *On understanding word order asymmetries*, 28 Theoretical Ling. 159, 164 (2002). This should be an incredibly uncontroversial point. The adjective "long" means one thing when modifying "story" and something else entirely when modifying "table." It is the noun that indicates which meaning should prevail.

This is true even when the noun is a term of art. For example, in common parlance, the word "hit" can take multiple meanings: It can mean "an impact or collision;" or "a successful stroke, performance, or production;" or a "dose of a narcotic drug."[22] But in baseball, "hit" has a technical meaning that refers specifically to instances "when a batter strikes the baseball into fair territory and reaches base without doing so on an error

---

[22] Hit, dictionary.com (last visited Nov. 16, 2018).

or a fielder's choice."[23] When discussing baseball, any adjective that modifies "hit" (i.e. *hard, long, powerful, magical, ugly*) must be understood in reference to that technical meaning, even if the full phrase might be understood differently in another context. Put another way, an adjective can hardly ever decontexualize a noun.

But that is the exact opposite of what the Ninth Circuit held. Despite the fact that in the context of federal litigation, "costs" is a "term of art" with a well-defined meaning, the Ninth Circuit concluded that by adding the word "full," Congress intended to strip "costs" of its technical meaning because all other interpretations would "effectively read[] the word 'full' out of the statute." That violates the basic grammatical principle that nouns govern adjectives, not the other way around. It is as absurd as saying that a newspaper article reporting that "Steve Pearce had two full hits during Game 5 of the World Series," was accusing the Red Sox' MVP of taking drugs because keeping the technical meaning of "hit" would "effectively read[] the word 'full' out of the" sentence.

This is not just an exercise in abstract, grammatical theory. Our corpus research shows that the meaning of the word "full" is always determined in reference to the word it is modifying. Consider the following lists of common nouns paired with "full" in each of our three main corpora[24]

---

[23] Hit (H), Major League Baseball, http://m.mlb.com/glossary/standard-stats/hit (last visited Nov. 16, 2018).

[24] These lists were generated using a corpus tool known as collocation. Collocation is the tendency for words to co-occur with

> COCA: Full time, full moon, full range, full day, full name, full circle, full force, full swing, full speed, full house, full length, full lips, full spectrum, full cost, full Senate, full scholarship, full benefits, full size, full accounting, full value, full screen, full frontal, full count, full beard, full gallop.
>
> COHA (1826-1836): Full extent, full speed, full view, full length, full assurance, full effect, full confidence, full force, full share, full moon, full power, full justice, full knowledge, full possession, full gallop, full enjoyment, full tide, full heart, full hour, full credit, full development, full value, full operation, full stop.
>
> Corpus of the Supreme Court (1826-1836): full consideration, full operation, full examination, full dominion, full satisfaction, full execution, full compensation, full investigation, full conviction, full possession, full description, full discussion, full exposition, full administration, full confirmation, full contemplation, full contribution, full exhibition, full illustration, full indemnification, full justification, full reinvestigation, full recognition, full remuneration.

As Petitioners suggest, the word "full" often "denotes that the limits of whatever is being modified have been reached." (Pet. Br. 16). When that is true, those outer limits (not to mention the unit of measurement) are determined by the noun *in context*. This point can be

---

other words in the corpus. You can recreate our search by searching for noun collocates one place to the right of "full" in each corpus. To access the full lists, see the Collocates-Full tab of https://goo.gl/Tmjnra

demonstrated with a quick review of concordance lines drawn from the COCA of the phrase "full house."

- The full House vote is expected tomorrow.
- Expect a full house for Ben Folds.
- James also returns to a full house: last year he married his Danish fiancé[e], Louise Holm, and the couple has a fifteen-month old daughter named Amelia.

In each case, the noun's fixed, contextualized meaning determines the scope and meaning of "full," not the other way around.[25] The addition of an adjective does not decontextualize the noun.

Furthermore, our research revealed many examples of "full" being used to describe a specific quantity or condition, even though that quantity or condition did not represent the literal maximum of whatever is being measured. Thus, a person who puts in a "full day of work" is presumed to have worked eight or nine hours, not twenty-four. Milk does not need to be anywhere close to 100% cream to be said to contain "full cream"—3-4% will suffice. Likewise, a parking lot can be considered "full" even if the handicap spots remain empty. In these cases, the government or some other external body actually plays a hand in defining "full".

In light of all this, the Ninth Circuit's interpretation is linguistically untenable. As explained below, the contextualized meaning of "costs" has been fixed by statute. The addition of the word "full" should

---

[25] We must distinguish these examples from instances where the addition of an adjective creates a compound noun, such as the term "full house" in poker.

not—indeed cannot—change that. Instead, the most natural way to read the statute is to interpret "full" as simply clarifying that a court may award all of the costs allowed under Section 1920 and no more.[26]

III. The word "full" is often delexicalized.

Another defect in the Ninth Circuit's opinion is that it is based on a clumsy application of the rule against surplusage. Although it is true that "[i]f possible, every word and every provision [of a statute] is to be given *effect*," Antonin Scalia & Bryan A. Garner, Reading Law 174 (2012) (emphasis added), this does not mean that each word must make some unique definitional contribution to the sentence's meaning. Words can often play a functional role. As a "linguistic canon of construction" the rule against surplusage is intended to merely "reflect[] the nature or use of language generally." Francis Bennion, Statutory Interpretation 805 (2d ed. 1992). In other words, the applicability of this canon in a given context "stand[s] or fall[s] by [its] accuracy in reflecting relevant linguistic practices" of the relevant speech community, whether that be the general population or "those who write and read legislation." William Baude & Stephen E. Sachs, *The Law of Interpretation*, 130 Harv. L. Rev. 1079, 1125 (2017).

---

[26] This reading is strengthened further by our corpus analysis, which reveals that one of the most common adjectives to modify *costs* in Supreme Court opinions is "double." See, *e.g.*, *Chambers* v. *NASCO, Inc.*, 501 U.S. 32, 40 (1991) ("The court imposed appellate sanctions in the form of attorney's fees and double costs."); *Talamini* v. *Allstate Ins. Co.*, 470 U.S. 1067, 1068 (1985) ("[A]ppellee requests the Court to award it 'double costs.'") An award of double costs references costs as defined in statute, not double all litigation expenses. In § 505, though, Congress limited recovery to just "*full* costs."

A. "Delexicalization" is common in American English.

Linguists have long noted that individual "words in English do not normally constitute independent selections." John Sinclair, Trust the Text: Language, Corpus and Discourse 20 (2004). Instead, we think, speak, and write in linguistic bundles. "Rather than choosing each word carefully and independently to convey an intended meaning, the choice of a given word often conditions the choice of the next word." Stephen Mouritsen, *Contract Interpretation with Corpus Linguistics*, 94 Wash. L. Rev. (forthcoming 2019). As a result, "[t]he meaning of words chosen together is different from their independent meanings." Sinclair, Trust the Text 19.. Linguists refer to this phenomenon as "co-selection." Mouritsen, 94 Wash. L. Rev.

Consequently, linguists have observed "a broad general tendency for frequent words"—such as "full", the 504th most common word in modern English[27]—"to have less of a clear and independent meaning than less frequent words." John Sinclair, Corpus, Concordance, Collocation (1991). This phenomenon is known as "delexicalization" and results in a "reduction of the *distinctive contribution made by that word*" to the overall meaning of any given sentence. *Id.* (emphasis added).

Delexicalized adjectives abound in English: "clenched fist," "free gift," "advanced planning," "full capacity," "past experience," etc.[28] "In all these cases if

---

[27] See Mark Davies, Word Frequency Data, https://www.wordfrequency.info/free.asp?s=y (last visited Nov. 16, 2018).

[28] For a more extensive list of common redundant expressions in English, see Richard Nordquist, *200 Common Redundancies in English*, ThoughtCo, https://www.thoughtco.com/common-redundancies-in-english-1692776 (last visited Nov. 15, 2018).

the adjective is removed there is no difficulty whatsoever in interpreting the meaning of the noun in exactly the way it was intended." Sinclair, Trust the Text at 21. Although such adjectives add little or nothing to the overall *meaning* of the sentence, this does not mean that they have no "effect." Scalia at 174; see also Sinclair, Trust the Text, at 22 ("The adjective ... is simply underlining part of the meaning of the noun."); Michael Guest, *Which Words?* 22 Japan Ass'n for Lang. Teaching 169, 171 (2000) (an adjective that has been delexicalized "has a use or function rather than a meaning"); Elena Tognini-Bonelli, Corpus Linguistics at Work 116 (2001) (delexicalized words "acquire[] other functions within a larger unit"). They primarily serve an "emphatic descriptive modifying function," V. Gonzalez-Diaz, *Great Big Stories and Tiny Little Changes: Tautological size-adjective clusters in present-day English*, 22 Eng. Lang. & Ling. 499 (2018), a way of drawing attention to or "underlining part of the meaning of the noun" Sinclair, Trust the Text at 22.

Indeed, linguist John Sinclair studied the delexicalization of "full" nearly thirty years ago. *Id.*at 22. Our analysis confirm his findings. We found that often—in both ordinary American English and legalese—"full" is not essential to the overall semantic meaning of a sentence, but instead functions as a "type[] of reassurance more than anything else." *Id.* Accordingly, the Ninth Circuit's invocation of the surplusage canon distorts rather than reflects the true nature and use of the word "full".

B. "Full" is often delexicalized in ordinary American English.

Just as the word "clenched" only emphasizes the inherent nature of a "fist," the word "full" often serves to

emphasize the already-presumed completeness of the noun it modifies. Our analysis revealed that "full" was delexicalized in this manner in approximately 46% of all of the concordance lines extracted from the COHA.[29] But our sample included a full 152 instances of the expression "full *of*" (i.e. full of hope, full of fear, full of cherries)—many of which were abstract and metaphorical. When one excludes these "full of" concordance lines and focuses exclusively on instances where "full" was directly modifying a noun, the number of delexicalized sentences goes up to 64%.

More important than the overall statistics is the fact that the sentences in which "full" was delexicalized were those that most resembled the usage and structure of Section 505—sentences where "full" immediately preceded the noun and was not part of a broader idiom. Consider the following real-world examples that appeared in actual texts published during the decade the Copyright Act was first passed. We've placed the word "full" placed in brackets to more clearly show that its absence would not change meaning:

- Clanton's blunder in this particular exposed the [full] depth of his villainy[.][30]

---

[29] Our results can be accessed at https://goo.gl/Tmjnra under the "COHA-Full" tab. To replicate our results, search COHA for "full" for the years 1826-1836 and generate a simple random sample of 1000. Corpus of Historical American English, BYU Corpora, https://corpus.byu.edu/coha/.

[30] H.R. Howard, The History of Virgil A. Stewart and his Adventure in capturing and Exposing the Great 'Western Pirate' and his Gang 165 (1836).

- She has the [full] power, as a man of [full] age has, to make any contract concerning property.[31]
- His ancestors were of the number of those Protestants, who were driven from France on the occasion of the revocation of the edict of Nantz, and bore their [full] share in the sufferings of that disastrous time.[32]
- Even while Juan doubted not that Guzman's skill and fortitude would insure him a [full] triumph, and final liberation, he could not but be struck with horror[.][33]
- Being, however, a [full] believer in the Malthusian theory, that population is always disposed to increase so rapidly as to be threatened with starvation ... he asserts that population has increased much faster than capital[.][34]

The same is true for the sample drawn from the COCA, where "full" appeared to be delexicalized 36% of the time. Again, note the grammatical similarities of these examples to the statutory provision at issue here:

---

[31] William Sullivan & George Barrell Emerson, The Political Class Book: Intended to Instruct the Higher Classes in Schools 120 (1831).

[32] William Jay, *The Life of John Jay*, 37 N. Am. Rev. 315, 316 (1833).

[33] Robert Montgomery Bird, The Infidel's Doom 109 (1840).

[34] Henry Charles Carey, Essay on the Rate of Wages, with an Examination of the Causses of the Differences in the Condition of the Labouring Population Throughout the World 232 (1835).

- That opportunity will not last, if we allow him to find the [full] protection of the barricades in the city.
- Treatment begins with a thorough sexual history to assess the [full] extent of the problem.
- Then he smiles and kisses me [full] on the mouth.
- That amount of data—taking [full] advantage of current communication techniques—can be transmitted in a second.
- Among their provisions, the new standards require tank car steel to be a [full] inch thick.

The rate of delexicalization rose once again—this time by a full fourteen percentage points—when the "full of" constructions were excluded.

Oracle's own writing epitomizes this trend. Consider the following actual examples taken from Oracle's website:

- JavaOne4Kids provides a [full] day of learning opportunities for students.
- Read the [full] press release from OpenWorld 2017.
- The report emphasizes the importance of incorporating AI into an overall business strategy to realize its [full] potential for guests, employees and the organization as a whole.
- HR and Finance: Unlocking the [Full] Potential of the Strategic CRHO.
- You can read the [full] details of the adoption story on Adam's blog.

Of all our corpora, Oracle's website was the most likely to use "full" in a way that resembled its use in Section 505, resulting in an overall delexicalization rate of 78%.[35]

C. "Full" is delexicalized in legalese.

Delexicalization appears to be an even more common phenomenon in legal writing. Our analysis of 19th Century opinions of this Court revealed that during the decade in which the relevant language was first passed (1826-1836), "full" was delexicalized just under 70% of the time. Unsurprisingly, these concordance lines looked a lot like provision at issue here. The following are just a few of hundreds of examples drawn from the corpus in which the effect of "full" was principally one of emphasis rather than additional meaning.

- He has not said ... that in every possible case, a fraudulent intent on the part of the grantor would avoid a deed to a bona fide purchaser for a [full] and valuable consideration having no knowledge of the fraud.[36]
- The principle asserted is that the creditor has a right to his debtor 's property by virtue of the obligation of the contract, to the [full] satisfaction of the debt[.][37]
- A monition from the Admiralty was sued out to the captor's agent, to respond to the captain's

---

[35] Our raw data can be accessed under the "Oracle Contracts" tab at https://goo.gl/Tmjnra.

[36] *Brooks* v. *Marbury*, 24 U.S. 78 (1826).

[37] *Ogden* v. *Saunders*, 25 U.S. 213, 310 (1827).

> demand for freight, to the [full] amount decreed to him.[38]
>
> - The judgment was in [full] force, and warranted the issuing of this execution[.][39]

The same is true in modern legal writing, as exemplified by Oracle's own contracts. There "full" was delexicalized 70.6% of the time, as in the following examples:

- [T]he Borrower shall timely pay the [full] amount deducted to the relevant Government Authority in accordance with applicable law.
- Tenant has designated Richard Henson as its sole representative with respect to the matters set forth in this Work Letter, who shall have [full] authority and responsibility to act on behalf of the Tenant as required in this Work Letter.
- Epstein agrees that the foregoing payment constitutes consideration for [full], complete and final settlement and release of any and all claims[.]
- Except as provided herein, all other terms and condition of the Lease remain in [full] force and effect.

These results should not surprise anyone familiar with legal drafting. As Professor Linda D. Jellum put it, "Legal drafters often include redundant language on purpose to cover any unforeseen gaps or simply for no good reason at all. And legislators are not likely to

---

[38] *Ramsay* v. *Allegre*, 25 U.S. 611, 630 (1827).

[39] *Bank of United States* v. *Bank of Wash.*, 31 U.S. 8 (1832).

waste time or energy arguing to remove redundancy when there are more important issues to address." Linda D. Jellum, Mastering Statutory Interpretation 104 (2008). Even Justice Scalia—a champion of the rule against surplusage—acknowledged that "[s]ometimes drafters *do* repeat themselves and *do* include words that add nothing of substance, either out of a flawed sense of style or to engage in the ill-conceived but lamentably common belt-and-suspenders approach." Antonin Scalia & Bryan A. Garner, Reading Law 179 (2012). These observations were confirmed by a recent survey of Congressional staffers, which revealed that a large majority of legislative drafters intentionally "err on the side of redundancy to 'capture the universe' or 'because [they] want to be sure [they] hit it.'" Abbe R. Gluck & Lisa Schultz Bressman, *Statutory Interpretation From the Inside—An Empirical Study of Congressional Drafting, Delegation, and the Canons: Part I*, 65 Stan. L. Rev. 901, 934 (2013).

In light of this, the most natural reading of Section 505 is to interpret "full" as a delexicalized adjective. It's "effect" is merely to underline the fact that all of the costs available under Section 1920 may be awarded, rather than to alter the ordinary legal meaning of "costs."

IV. Judges and Congress consistently use the phrase "full costs" in a way that excludes expert witness fees.

Moreover, as Petitioners have explained, both Congress and the Supreme Court have "long distinguished costs, fees, and expenses." (Pet. Br. 26 (internal quotations omitted)). Our findings discussed below confirm that this true. This strengthens the reliability of our corpus analysis.

1. As Professors Larry Solan and Tammy Gales put it, "the strongest cases for using corpus analysis are ones in which not only does one meaning predominate over an alternative meaning in an appropriate corpus, but the second, less common meaning is generally expressed using language *other than* the language in the disputed statute." Lawrence M. Solan & Tammy Gales, *Corpus Linguistics as a Tool in Legal Interpretation*, 2017 B.Y.U. L. Rev. 1311, 1315 (emphasis added).

The Court's approach in *West Virginia University Hospitals, Inc.* v. *Casey* is informative. 499 U.S. 83 (1991). In *Casey*, the Court held that expert witness fees were not considered part of attorneys' fees, surveying federal case law to show that judges typically referred to expert witness fees and attorneys' fees as "separate categories of expense." *Id.* at 93-95 (listing dozens of cases). Thus, the fact that the Seventh Circuit once stated that the "costs awarded by the [district] court" included both "$1,700 for plaintiff's expert witness ... *and* attorneys' fees $15,660" was linguistic evidence that it did not consider expert witness fees to be "an *element of* attorney's fees." *Id.* at 92-93.The Court listed thirty-four such cases, while noting that it could find "no support" for the inverse proposition. *Id.* at 94.

2. The same line of reasoning can be used to assess whether judges consider expert witness fees and other nontaxable costs to be *elements of* "full costs" or instead treat them as "separate categories of expense." Modern legal databases make this inquiry much easier and systematic than it was in 1991. A simple search of all federal cases in Westlaw for the phrase "full costs *and*" reveals hundreds of cases where judges have

clearly felt the phrase "full costs" did not include certain categories of litigation expenses.

Attorneys' fees is the most common category excluded[40]—with judges specifically contrasting attorneys' fees (or its synonyms) with "full costs" in 392 of the 486 cases. Other terms specifically excluded include "expenses,"[41] "fees,"[42] "expert witness fees,"[43]

---

[40] See, *e.g.*, *United States* v. *Reese*, 92 U.S. 214 (2 Otto) (1875) ("[F]or every offence, [he] shall forfeit and pay the sum of $500 to the person aggrieved thereby, to be recovered by an action on the case with full costs, and such allowance for counsel-fees as the court shall deem just."); *J&J Sports Prods., Inc.* v. *Evolution Entertainment Grp., LLC*, 2014 WL 2066809 (E.D. La. 2014) ("The statute allows an aggrieved party who prevails to recover full costs and reasonable attorney's fees.").

[41] The category of "expenses" was excluded from "full costs" in 40 cases. See, *e.g. Waste Control Specialists, LLC* v. *Envirocare of Tex., Inc.*, 199 F.3d 781, 787 (5th Cir. 2000) (opinion withdrawn) ("We therefore conclude that WCS is entitled under the statute to the full costs *and expenses*[.]); *In re Lipset*, 119 F. 379 (S.D.N.Y. 1902) ("[T]he court will speedily find a remedy therefor[e] by the imposing of the full costs *and expenses* occasioned by such undue exercise of the power of examination.");

[42] Contrasted with "full costs" in 9 cases. See, *e.g. Innovative Sports Management, Inc.* v. *Adriana, Inc.*, 2014 WL 1491339, at *4 (E.D. Va. Feb. 3, 2014) ("Plaintiff's Complaint sought ... full costs and expenses."); *Imaginary Images Inc.* v. *Evans*, 2009 WL 2488004, at *2 (E.D. Va. Aug. 12, 2009) ("Defendants make no specific argument as to the validity of these costs ... and therefore [are] not entitled to full costs and fees.");

[43] Contrasted with "full costs" in 5 cases. See, *e.g.*, *Express LLC* v. Forever 21, 2010 WL 11512410 (C.D. Cal. Nov. 15, 2010) ("[T]he Copyright Act provision permitting recovery of "full costs" ... makes no reference to witness fees.").

"professional fees,"[44] and "interest."[45] Some of these cases are more than two centuries old, predating the passage of the relevant language in the Copyright Act. For example, in *The Betsy*, Justice Story ordered the restoration of captured cargo "upon the payment of the full costs and expenses of the captors."[46]

This widespread practice—across time and a variety of diverse jurisdictions—of excluding various categories of litigation expenses from "full costs" forecloses the holding below. Were "full costs" to mean all litigation expenses as the Ninth Circuit has held, there would have been no need for these judges to specifically mention attorneys' fees, expert witness fees, expenses, and the like. "Full costs" would have covered it all.

Of course, this is not dispositive. As mentioned above, legal language is rife with duplicative language. But it would be odd to dismiss this evidence of linguistic practice within the judiciary as merely duplicative language in order to support a holding premised on a presumption that legal language is *never* duplicative—especially when corpus evidence suggests at least two more natural readings of the statute.

---

[44] *SEC* v. *Utsick*, 2009 WL 1606511, at *9 ("The SEC's report recommended that Michaelson recover full costs and $2,643.75 in fees for certified public accountant['s] ... work. With respect to the remaining $41,050 in professional fees, the SEC recommended that Michaelson submit biographies and descriptions of the work performed by other employees.).

[45] Contrasted with "full costs" in 6 cases. See, *e.g.*, *Hygrade No. 24, the* v. *the Dynamic*, 143 F. Supp. 634 (E.D.N.Y. 1955) ("The libellant claims full costs and interest.").

[46] *The Betsy*, 3 F. Cas. 299, 303 (D. Mass. 1815) (Story, J.).

3. An analysis of Congressional use of the phrase "expert witness fees" specifically—the largest nontaxable cost at issue here—further supports this conclusion. "Expert witness fees" are mentioned in the U.S. Code eighty-two different times across sixty-one sections. In a majority of these (~54%) expert witness fees are lumped in with other "costs of litigation."[47]

But when this happens, it is almost always as a parenthetical using a particular formulation: "The court ... may award costs of litigation (including reasonable attorney and expert witness fees)."[48] This formulation is itself striking because it indicates that attorneys' fees and expert witness fees are typically not considered part of the "costs of litigation," or else they would not need to be singled out as included.[49] In

---

[47] Although "costs of litigation" is the most common formulation, sometimes they are called "costs of suit," *e.g.* 15 U.S.C. 2060(c); "costs of defending the action," *e.g.* 42 U.S.C. 9607(p)(7); "costs of filings and pursuing the protest," 31 U.S.C. 3554(c)(1)(A), etc.

[48] See, *e.g.*, 12 U.S.C. 5567(c)(5)(C); 49 U.S.C. 30171(b)(6)(B). Less often, Congress will use the exact same language but without the parentheses. See, *e.g.*, 30 U.S.C. 1427(c) ("The court ... may award costs of litigation, including attorney and expert witness fees).

[49] In a manner, it would be odd for someone say "The entire Supreme Court was there (including Justice Kagan)" unless there was some *reason* for the reader to otherwise expect Justice Kagan to be absent. But it would not be odd for someone to say "The entire Supreme Court was there (including Justice O'Connor)" since retired justices are not typically considered part of the Supreme Court.

fact, in the U.S. Code the term "costs of litigation" appears on its own three times as often as it does with these sorts of parenthetical provisos.[50]

Elsewhere in the U.S. Code, "expert witness fees" are categorized as "expenses" (19.3%), "damages" (2.2%), and "special damages" (2.2%).[51] Here too, expert witness fees are mentioned only parenthetically, implying that expert witness fees are only available when Congress specifically makes them available. Congress has not done so here.

---

[50] *Compare* Westlaw search for "costs of litigation" in U.S. Code (157) *with* "costs of litigation including" in U.S. Code (39).

[51] See, *e.g.* 5 U.S.C. 3330c(b) ("A prefence eligible who prevails ... shall be awarded ... expertwitness fees, and other litigation expenses."); 30 U.S.C. 1270(f) ("Any person who is injured ... may bring an action for damages (including ... expert witness fees"); 7 U.S.C. 26(h)(1)(C) ("Relief for an individual prevailing ... shall include ... compensation for any special damages sustained ... including ... expert witness fees). Our full dataset can be accessed under the "U.S. Code: Expert Witness Fees" tab at https://goo.gl/Tmjnra.

CONCLUSION

The Ninth Circuit's interpretation of "full" is linguistically untenable and therefore legally insufficient to alter the statutory definition of "costs." The decision should be reversed.

Respectfully submitted,

JAMES HEILPERN
Law & Corpus Linguistics Fellow
BYU LAW SCHOOL
415 JRCB
Provo, UT 84602

GENE C. SCHAERR
Counsel of Record
MICHAEL T. WORLEY
SCHAERR | JAFFE LLP
1717 K Street NW, Suite 900
Washington, DC 20006
(202) 787-1060
gschaerr@schaerr-jaffe.com

November 20, 2018

APPENDIX

APPENDIX: List of *Amici*

Laurence Anthony is a Professor in the Faculty of Science and Engineering at Waseda University, Japan.

William G. Eggington is the Ludwig-Weber-Siebach Humanities Professor in the Linguistics and English Department of Brigham Young University.

Tammy Gales is an Associate Professor of Comparative Literature, Languages, and Linguistics at Hofstra University.

Tim Grant is a Professor in Forensic Linguistics, Director of the Centre for Forensic Linguistics, and Associate Dean of Partnerships and Community Engagement at Aston University.

Stefan Th. Gries is a Professor of Linguistics at the University of California, Santa Barbara.

Benjamin Lee is the Corpus Linguistics Research Fellow at the J. Reuben Clark Law School at Brigham Young University.

Tony McEnry is a Distinguished Professor of English Language and Linguistics at Lancaster University.

Jeffrey R. Parker is an Assistant Professor of Linguistics & English Language at Brigham Young University.

Rui Sousa-Silva is an Assistant Professor of the Faculty of Arts and Post-doctoral Researcher at the Linguistics Centre of the University of Porto.

Lawrence Solan is the Don Forchelli Professor of Law and Director of Graduate Education at Brooklyn Law School.

Sara White is the former Corpus Linguistics Research Fellow at the J. Reuben Clark Law School at Brigham Young University.